IMAGE PROCESSING USING MEASURES OF SIMILARITY

Prior Applications

[0001] This application claims the benefit of U.S. Provisional Patent Application Serial Number 60/353,978, filed January 31, 2002, the contents of which are hereby incorporated by reference.

Government Rights

[0002] This invention was made with government support under Grant No.1-R44-CA-91618-01 awarded by the U.S. Department of Health and Human Services. The government has certain rights in the invention.

Field of the Invention

[0003] This invention relates generally to image processing. More particularly, in certain embodiments, the invention relates to segmentation of a sequence of colposcopic images based on measures of similarity.

Background of the Invention

[0004] It is common in the medical field to perform visual examination to diagnose disease. For example, visual examination of the cervix can discern areas where there is a suspicion of pathology. However, direct visual observation alone is often inadequate for identification of abnormalities in a tissue.

[0005] In some instances, when tissues of the cervix are examined *in vivo*, chemical agents such as acetic acid are applied to enhance the differences in appearance between

normal and pathological areas. Aceto-whitening techniques may aid a colposcopist in the determination of areas where there is a suspicion of pathology.

[0006] However, colposcopic techniques generally require analysis by a highly trained physician. Colposcopic images may contain complex and confusing patterns. In colposcopic techniques such as aceto-whitening, analysis of a still image does not capture the patterns of change in the appearance of tissue following application of a chemical agent. These patterns of change may be complex and difficult to analyze. Current automated image analysis methods do not allow the capture of the dynamic information available in various colposcopic techniques.

[0007] Traditional image analysis methods include segmentation of individual images. Segmentation is a morphological technique that splits an image into different regions according to one or more pre-defined criteria. For example, an image may be divided into regions of similar intensity. It may therefore be possible to determine which sections of a single image have an intensity within a given range. If a given range of intensity indicates suspicion of pathology, the segmentation may be used as part of a diagnostic technique to determine which regions of an image may indicate diseased tissue.

[0008] However, standard segmentation techniques do not take into account dynamic information, such as a change of intensity over time. This kind of dynamic information is important to consider in various diagnostic techniques such as aceto-whitening colposcopy. A critical factor in discriminating between healthy and diseased tissue may be the manner in which the tissue behaves throughout a diagnostic test, not just at a given time. For example, the rate at which a tissue whitens upon application of a chemical

agent may be indicative of disease. Traditional segmentation techniques do not take into account time-dependent behavior, such as rate of whitening.

Summary of the Invention

[0009] The invention provides methods for relating aspects of a plurality of images of a tissue in order to obtain diagnostic information about the tissue. In particular, the invention provides methods for image segmentation across a plurality of images instead of only one image at a time. In a sense, inventive methods enable the compression of a large amount of pertinent information from a sequence of images into a single frame. An important application of methods of the invention is the analysis of a sequence of images of biological tissue in which an agent has been applied to the tissue in order to change its optical properties in a way that is indicative of the physiological state of the tissue. Diagnostic tests which have traditionally required analysis by trained medical personnel may be automatically analyzed using these methods. The invention may be used, for example, in addition to or in place of traditional analysis.

[0010] The invention provides methods of performing image segmentation using information from a sequence of images, not just from one image at a time. This is important because it allows the incorporation of time effects in image segmentation. For example, according to an embodiment of the invention, an area depicted in a sequence of images is divided into regions based on a measure of similarity of the changes those regions undergo throughout the sequence. In this way, inventive segmentation methods incorporate more information and can be more helpful, for example, in determining a characteristic of a tissue, than segmentation performed using one image at a time. The phrases “segmentation of an image” and “segmenting an image,” as used herein, may

apply, for example, to dividing an image into different regions, or dividing into different regions an area in space depicted in one or more images (an image plane).

[0011] Segmentation methods of this invention allow, for example, the automated analysis of a sequence of images using complex criteria for determining a disease state which may be difficult or impossible for a human analyst to perceive by simply viewing the sequence. The invention also allows the very development of these kinds of complex criteria for determining a disease state by permitting the relation of complex behaviors of tissue samples during dynamic diagnostic tests to the known disease states of the tissue samples. Criteria may be developed using the inventive methods described herein to analyze sequences of images for dynamic diagnostic tests that are not yet in existence.

[0012] One way to relate a plurality of images to each other according to the invention is to create or use a segmentation mask that represents an image plane divided into regions that exhibit similar behavior throughout a test sequence. Another way to relate images is by creating or using graphs or other means of data representation which show mean signal intensities throughout each of a plurality of segmented regions as a function of time. Relating images may also be performed by identifying any particular area represented in an image sequence which satisfies given criteria.

[0013] In one aspect, the invention is directed to a method of relating a plurality of images of a tissue. The method includes the steps of: obtaining a plurality of images of a tissue; determining a relationship between two or more regions in each of two or more of the images; segmenting at least a subset of the two or more images based on the relationship; and relating two or more images of the subset of images based at least in part on the segmentation.

[0014] According to one embodiment, the step of obtaining a plurality of images of a tissue includes collecting an optical signal. In one embodiment, the optical signal includes fluorescence illumination from the tissue. In one embodiment, the optical signal includes reflectance, or backscatter, illumination from the tissue. In one embodiment, the tissue is illuminated by a white light source, a UV light source, or both. According to one embodiment, the step of obtaining images includes recording visual images of the tissue.

[0015] According to one embodiment, the tissue is or includes cervical tissue. In another embodiment, the tissue is one of the group consisting of epithelial tissue, colorectal tissue, skin, and uterine tissue. In one embodiment, the plurality of images being related are sequential images. In one embodiment, the chemical agent is applied to change its optical properties in a way that is indicative of the physiological state of the tissue. According to one embodiment, a chemical agent is applied to the tissue. In one embodiment, the chemical agent is selected from the group consisting of acetic acid, formic acid, propionic acid, butyric acid, Lugol's iodine, Shiller's iodine, methylene blue, toluidine blue, osmotic agents, ionic agents, and indigo carmine. In certain embodiments, the method includes filtering two or more of the images. In one embodiment, the method includes applying either or both of a temporal filter, such as a morphological filter, and a spatial filter, such as a diffusion filter.

[0016] In one embodiment, the step of determining a relationship between two or more regions in each of two or more of the images includes determining a measure of similarity between at least two of the two or more images. In one embodiment, determining the measure of similarity includes computing an N-dimensional dot product

of the mean signal intensities of two of the two or more regions. In one embodiment, the two regions (of the two or more regions) are neighboring regions.

[0017] According to one embodiment, the step of relating images based on the segmentation includes determining a segmentation mask of an image plane, where two or more regions of the image plane are differentiated. In one embodiment, the step of relating images based on the segmentation includes defining one or more data series representing a characteristic of one or more associated segmented regions of the image plane. In one embodiment, this characteristic is mean signal intensity.

[0018] According to one embodiment, the step of relating images includes creating or using a segmentation mask that represents an image plane divided into regions that exhibit similar behavior throughout the plurality of images. In one embodiment, the step of relating images includes creating or using graphs or other means of data representation which show mean signal intensities throughout each of a plurality of segmented regions as a function of time. In one embodiment, the step of relating images includes identifying a particular area represented in the image sequence which satisfies given criteria.

[0019] In another aspect, the invention is directed to a method of relating a plurality of images of a tissue, where the method includes the steps of: obtaining a plurality of images of a tissue; determining a measure of similarity between two or more regions in each of two or more of the images; and relating at least a subset of the images based at least in part on the measure of similarity. In one embodiment, the step of determining a measure of similarity includes computing an N-dimensional dot product of the mean signal

intensities of two of the two or more regions. In one embodiment, the two regions are neighboring regions.

[0020] In another aspect, the invention is directed to a method of determining a tissue characteristic. The method includes the steps of: obtaining a plurality of images of a tissue; determining a relationship between two or more regions in each of two or more of the images; segmenting at least a subset of the two or more images based at least in part on the relationship; and determining a characteristic of the tissue based at least in part on the segmentation.

[0021] According to one embodiment, the step of obtaining a plurality of images of a tissue includes collecting an optical signal. In one embodiment, the optical signal includes fluorescence illumination from the tissue. In one embodiment, the optical signal includes reflectance, or backscatter, illumination from the tissue. In one embodiment, the tissue is illuminated by a white light source, a UV light source, or both. According to one embodiment, the step of obtaining images includes recording visual images of the tissue.

[0022] According to one embodiment, the tissue is or includes cervical tissue. In another embodiment, the tissue is one of the group consisting of epithelial tissue, colorectal tissue, skin, and uterine tissue. In one embodiment, the plurality of images being related are sequential images. In one embodiment, the chemical agent is applied to change its optical properties in a way that is indicative of the physiological state of the tissue. According to one embodiment, a chemical agent is applied to the tissue. In one embodiment, the chemical agent is selected from the group consisting of acetic acid, formic acid, propionic acid, butyric acid, Lugol's iodine, Shiller's iodine, methylene blue, toluidine blue, osmotic agents, ionic agents, and indigo carmine. . In certain

embodiments, the method includes filtering two or more of the images. In one embodiment, the method includes applying either or both of a temporal filter, such as a morphological filter, and a spatial filter, such as a diffusion filter. In one embodiment, the method includes processing two or more images to compensate for a relative motion between the tissue and a detection device.

[0023] According to one embodiment, the step of determining a relationship between two or more regions in each of two or more of the images includes determining a measure of similarity between at least two of the two or more images. In certain embodiments, determining this measure of similarity includes computing an N-dimensional dot product of the mean signal intensities of two of the two or more regions. In one embodiment, the two regions are neighboring regions.

[0024] According to one embodiment, the segmenting step includes analyzing an aceto-whitening signal. In one embodiment, the segmenting step includes analyzing a variance signal. In one embodiment, the segmenting step includes determining a gradient image.

[0025] According to one embodiment, the method includes processing one or more optical signals based on the segmentation. In one embodiment, the method includes filtering at least one image based at least in part on the segmentation.

[0026] In certain embodiments, the step of determining a characteristic of the tissue includes determining one or more regions of the tissue where there is suspicion of pathology. In certain embodiments, the step of determining a characteristic of the tissue includes classifying a region of tissue as one of the following: normal squamous tissue, metaplasia, Cervical Intraepithelial Neoplasia, Grade I (CIN I), and Cervical Intraepithelial Neoplasia, Grade II or Grade III (CIN II/CIN III).

[0027] In another aspect, the invention is directed to a method of determining a characteristic of a tissue. The method includes the steps of: (a) for each of a first plurality of reference sequences of images of tissue having a first known characteristic, quantifying one or more features of each of a first plurality of mean signal intensity data series corresponding to segmented regions represented in each of the first plurality of reference sequences of images; (b) for a test sequence of images, quantifying one of more features of each of one or more mean signal intensity data series corresponding to one or more segmented regions represented in the test sequence of images; and (c) determining a characteristic of a tissue represented in the test sequence of images based at least in part on a comparison between the one or more features quantified in step (a) and the one or more features quantified in step (b).

[0028] According to one embodiment, step (c) includes repeating step (a) for each of a second plurality of reference sequences of images of tissue having a second known characteristic. In one embodiment, step (c) includes applying a classification rule based at least in part on the first plurality of reference sequences and the second plurality of reference sequences. In one embodiment, step (c) includes performing a linear discriminant analysis to determine the classification rule. In one embodiment, one of the one or more features of step (a) includes the slope of a curve at a given point fitted to one of the plurality of mean signal intensity data series. According to one embodiment, the method includes determining the segmented regions of the test sequence of images by analyzing an acetowhitening signal. In one embodiment, the first known characteristic is CIN II/CIN III and the second known characteristic is absence of CIN II/CIN III.

Brief Description of the Drawings

[0029] The objects and features of the invention can be better understood with reference to the drawings described below, and the claims. The drawings are not necessarily to scale, emphasis instead generally being placed upon illustrating the principles of the invention. In the drawings, like numerals are used to indicate like parts throughout the various views.

[0030] The patent or application file contains at least one drawing executed in color. Copies of this patent or patent application publication with color drawing(s) will be provided by the U.S. Patent and Trademark Office upon request and payment of the necessary fee.

[0031] Figure 1 is a schematic flow diagram depicting steps in the analysis of a sequence of images of tissue according to an illustrative embodiment of the invention.

[0032] Figure 2A depicts human cervix tissue and shows an area of which a sequence of images are to be obtained according to an illustrative embodiment of the invention.

[0033] Figure 2B depicts the characterization of a discrete signal from a sequence of images of tissue according to an illustrative embodiment of the invention.

[0034] Figure 3 shows a series of graphs depicting mean signal intensity of a region as a function of time, as determined from a sequence of images according to an illustrative embodiment of the invention.

[0035] Figure 4A depicts a "maximum" RGB image representation used in preprocessing a sequence of images of tissue according to an illustrative embodiment of the invention.

[0036] Figure 4B depicts the image representation of Figure 4A after applying a manual mask, used in preprocessing a sequence of images of tissue according to an illustrative embodiment of the invention.

[0037] Figure 4C depicts the image representation of Figure 4B after accounting for glare, used in preprocessing a sequence of images of tissue according to an illustrative embodiment of the invention.

[0038] Figure 4D depicts the image representation of Figure 4C after accounting for chromatic artifacts, used in preprocessing a sequence of images of tissue according to an illustrative embodiment of the invention.

[0039] Figure 5 shows a graph illustrating the determination of a measure of similarity of time series of mean signal intensity for each of two regions according to an illustrative embodiment of the invention.

[0040] Figure 6 is a schematic flow diagram depicting a region merging approach of segmentation according to an illustrative embodiment of the invention.

[0041] Figure 7A represents a segmentation mask produced using a region merging approach according to an illustrative embodiment of the invention.

[0042] Figure 7B shows a graph depicting mean signal intensities of segmented regions represented in Figure 7A as functions of time according to an illustrative embodiment of the invention.

[0043] Figure 8 is a schematic flow diagram depicting a robust region merging approach of segmentation according to an illustrative embodiment of the invention.

[0044] Figure 9A represents a segmentation mask produced using a robust region merging approach according to an illustrative embodiment of the invention.

[0045] Figure 9B shows a graph depicting mean variance signals of segmented regions represented in Figure 9A as functions of time according to an illustrative embodiment of the invention.

[0046] Figure 10 is a schematic flow diagram depicting a clustering approach of segmentation according to an illustrative embodiment of the invention.

[0047] Figure 11A represents a segmentation mask produced using a clustering approach according to an illustrative embodiment of the invention.

[0048] Figure 11B shows a graph depicting mean signal intensities of segmented regions represented in Figure 11A as functions of time according to an illustrative embodiment of the invention.

[0049] Figure 11C represents a segmentation mask produced using a clustering approach according to an illustrative embodiment of the invention.

[0050] Figure 11D shows a graph depicting mean signal intensities of segmented regions represented in Figure 11C as functions of time according to an illustrative embodiment of the invention.

[0051] Figure 12 is a schematic flow diagram depicting a watershed approach of segmentation according to an illustrative embodiment of the invention.

[0052] Figure 13 represents a gradient image used in a watershed approach of segmentation according to an illustrative embodiment of the invention.

[0053] Figure 14A represents a segmentation mask produced using a watershed approach according to an illustrative embodiment of the invention.

[0054] Figure 14B represents a segmentation mask produced using a watershed approach according to an illustrative embodiment of the invention.

[0055] Figure 15A represents a seed region superimposed on a reference image from a sequence of images, used in a region growing approach of segmentation according to an illustrative embodiment of the invention.

[0056] Figure 15B represents the completed growth of the "seed region" of Figure 15A using a region growing approach according to an illustrative embodiment of the invention.

[0057] Figure 16A represents a segmentation mask produced using a combined clustering approach and robust region merging approach according to an illustrative embodiment of the invention.

[0058] Figure 16B shows a graph depicting mean signal intensities of segmented regions represented in Figure 16A as functions of time according to an illustrative embodiment of the invention.

[0059] Figure 17A represents a segmentation mask produced using a combined clustering approach and watershed technique according to an illustrative embodiment of the invention.

[0060] Figure 17B shows a graph depicting mean signal intensities of segmented regions represented in Figure 17A as functions of time according to an illustrative embodiment of the invention.

[0061] Figure 18A represents a segmentation mask produced using a two-part clustering approach according to an illustrative embodiment of the invention.

[0062] Figure 18B shows a graph depicting mean signal intensities of segmented regions represented in Figure 18A as functions of time according to an illustrative embodiment of the invention.

[0063] Figure 19 depicts the human cervix tissue of Figure 2A with an overlay of manual doctor annotations made after viewing an image sequence.

[0064] Figure 20A is a representation of a segmentation mask produced using a combined clustering approach and robust region merging approach with a correspondingly-aligned overlay of the manual doctor annotations of Figure 19, according to an embodiment of the invention.

[0065] Figure 20B is a representation of a segmentation mask produced using a combined clustering approach and morphological technique with a correspondingly-aligned overlay of the manual doctor annotations of Figure 19, according to an embodiment of the invention.

[0066] Figure 21A depicts a reference image of cervical tissue of a patient from a sequence of images obtained during an acetowhitening test according to an illustrative embodiment of the invention.

[0067] Figure 21B depicts an image from the sequence of Figure 21A after applying a manual mask, accounting for glare, and accounting for chromatic artifacts according to an illustrative embodiment of the invention.

[0068] Figure 21C shows a graph depicting mean signal intensities of segmented regions for the sequence of Figure 21A determined using a clustering segmentation approach according to an illustrative embodiment of the invention.

[0069] Figure 21D represents a map of regions of tissue as segmented in Figure 21C classified as either high grade disease tissue or not high grade disease tissue using a classification algorithm according to an illustrative embodiment of the invention.

Description of the Illustrative Embodiment

[0070] In general, the invention provides methods for image segmentation across a plurality of images. Segmentation across a plurality of images provides a much more robust analysis than segmentation in a single image. Segmentation across multiple images according to the invention allows incorporation of a temporal element (e.g., the change of tissue over time in a sequence of images) in optics-based disease diagnosis. The invention provides means to analyze changes in tissue over time in response to a treatment. It also provides the ability to increase the resolution of segmented imaging by increasing the number of images over time. This allows an additional dimension to image-based tissue analysis, which leads to increase sensitivity and specificity of analysis. The following is a detailed description of a preferred embodiment of the invention.

[0071] The schematic flow diagram 100 of Figure 1 depicts steps in the analysis of a sequence of images of tissue according to an illustrative embodiment of the invention. Figure 1 also serves as a general outline of the contents of this description. Each of the steps of Figure 1 is discussed herein in detail. Briefly, the steps include obtaining a sequence of images of the tissue 102, preprocessing the images 104, determining a measure of similarity between regions in each of the images 108, segmenting the images 110, relating the images 112, and finally, determining a tissue characteristic 114. Though not pictured in Figure 1, the steps may be preceded by application of a chemical agent onto the tissue, for example. In other embodiments, a chemical agent is applied during the performance of the steps of the schematic flow diagram 100 Figure 1.

[0072] Among the key steps of the inventive embodiments discussed here are determining a measure of similarity between regions of tissue represented in a sequence of images and segmenting the images based on the measure of similarity. Much of the mathematical complexity presented in this description regards various methods of performing these key steps. As will become evident, different segmentation methods have different advantages. The segmentation techniques of the inventive embodiments discussed herein include region merging, robust region merging, clustering, watershed, and region growing techniques, as well as combinations of these techniques.

[0073] Figures 2A and 2B relate to step 102 of Figure 1, obtaining a sequence of images of the tissue. Although embodiments of the invention are not limited to aceto-whitening tests, an exemplary sequence of images from an aceto-whitening test performed on a patient is used herein to illustrate certain embodiments of the invention. Figure 2A depicts a full-frame image 202 of a human cervix after application of acetic acid, at the start of an aceto-whitening test. The inset image 204 depicts an area of interest to be analyzed herein using embodiment methods of the invention. This area of interest may be determined by a technician or may be determined in a semi-automated fashion using a multi-step segmentation approach such as one of those discussed herein below.

[0074] Figure 2B depicts the characterization 206 of a discrete signal $\mathbf{w}(i,j;t)$ from a sequence of images of tissue according to an illustrative embodiment of the invention. The signal could be any type of image signal of interest known in the art. In the illustrative embodiment, the signal is an intensity signal of an image.

[0075] In the illustrative embodiment, images of an area of interest are taken at N time steps $\{t_0, t_1, \ldots, t_{N-1}\}$. In one embodiment, time t_0 corresponds to the moment of

application of a chemical agent to the tissue, for instance, and time t_{N-1} corresponds to the end of the test. In another embodiment, time t_0 corresponds to a moment following the application of a chemical agent to the tissue. For example, let $\mathbb{I} = \{0,\ldots,r-1\} \times \{0,\ldots,c-1\}$ and $\mathbb{T} = \{n_0,\ldots,n_{N-1}\}$ be the image and time domains, respectively, where r is the number of rows and c is the number of columns. Then, $r \times c$ discrete signals $\mathbf{w}(i,j;t)$ may be constructed describing the evolution of some optically-detectable phenomena, such as aceto-whitening, with time. For an aceto-whitening example, the "whiteness" may be computed from RGB data of the images. There are any number of metrics which may be used to define "whiteness." For instance, an illustrative embodiment employs an intensity component, CCIR 601, as a measure of "whiteness" of any particular pixel, defined in terms of red (R), green (G), and blue (B) intensities as follows:

$$I = 0.299R + 0.587G + 0.114B \quad . \tag{1}$$

[0076] The "whitening" data is then given by $\mathbf{w}(i,j;t) = I(i,j;n)$, for example. Alternatively, the signal $\mathbf{w}(i,j;t)$ is defined in any of a multiplicity of other ways. The characterization 206 of Figure 2B shows that the intensity signal $\mathbf{w}(i,j;t)$ has a value corresponding to each discrete location (i,j) in each of the images taken at N discrete time steps. According to the illustrative embodiment, a location (i,j) in an image corresponds to a single pixel. In an aceto-whitening example, since it is the whitening of the cervix that is of interest and not the absolute intensity of the cervix surface, the whitening signals are background subtracted. In one example of background subtraction, each of the signals corresponding to a given location at a particular time step are transformed by subtracting the initial intensity signal at that location as shown in Equation (2):

$$\mathbf{w}(i,j;n) \mapsto \mathbf{w}(i,j;n) - \mathbf{w}(i,j;n_0),\ \forall n \in \mathrm{T} \quad . \tag{2}$$

[0077] Noise, glare, and sometimes chromatic artifacts may corrupt images in a sequence. Signal noise due to misaligned image pairs and local deformations of the tissue may be taken into account as well. Alignment functions and image restoration techniques often do not adequately reduce this type of noise. Therefore, it may be necessary to apply temporal and spatial filters.

[0078] Figure 3 relates to part of step 104 of Figure 1, preprocessing the images. Figure 3 shows a series of graphs depicting mean signal intensity 304 of a pixel as a function of time 306, as determined from a sequence of images according to an illustrative embodiment of the invention. The graphs depict application of a morphological filter, application of a diffusion filter, modification of intensity data to account for background intensity, and normalization of intensity data, according to an illustrative embodiment of the invention.

[0079] According to the illustrative embodiment, a first filter is applied to the time axis, individually for each pixel. The images are then spatially filtered. Graph 302 of Figure 3 depicts the application of both a temporal filter and a spatial filter at a representative pixel. The original data is connected by a series of line segments 308. It is evident from graph 302 that noise makes the signal choppy and adversely affects further analysis if not removed.

[0080] For temporal filtering, the illustrative embodiment of the invention applies the morphological filter of Equation (3):

$$\mathbf{w}(t) \odot b = \frac{1}{2}\left[(\mathbf{w} \circ b) \bullet b + (\mathbf{w} \bullet b) \circ b\right], \qquad (3)$$

where b is the *structuring element*, $\circ$ is the *opening* operator, and $\bullet$ is the *closing* operator. According to the illustrative embodiment, the structuring element has a half

circle shape. The temporally-filtered data is connected by a series of line segments 310 in the graph 302 of Figure 3. The noise is decreased from the series 308 to the series 310.

[0081] Illustratively, the images are then spatially filtered, for example, with either an isotropic or a Gaussian filter. A diffusion equation implemented by an illustrative isotropic filter may be expressed as Equation (4):

$$\frac{\partial \mathbf{w}(i,j)}{\partial \tau} = k\nabla \cdot \nabla \mathbf{w} = k\Delta \mathbf{w} \quad , \qquad (4)$$

where ∇ is the gradient operator, Δ is the Laplacian operator, and τ is the diffusion time (distinguished from the time component of the whitening signal itself). An isotropic filter is iterative, while a Gaussian filter is an infinite impulse response (IIR) filter. The iterative filter of Equation (4) is much faster than a Gaussian filter, since the iterative filter allows for increasing smoothness by performing successive iterations. The Gaussian filter requires re-applying a more complex filter to the original image for increasing degrees of filtration. According to the illustrative embodiment, the methods of the invention perform two iterations. However, in other embodiments, the method performs one iteration or three or more iterations. The spatially-filtered data for a representative pixel is connected by a series of line segments 312 in graph 302 of Figure 3. The noise is decreased from series 310 to series 312.

[0082] Graph 314 of Figure 3 shows the application of Equation (2), background subtracting the intensity signal 304. Graph 318 of Figure 3 shows the intensity signal data following normalization 320. In the illustrative embodiment, as explained below in further detail, normalization includes division of values of the intensity signal 304 by a reference value, such as the maximum intensity signal over the sequence of images. Glare and chromatic artifacts can affect selection of the maximum intensity signal; thus,

in an illustrative embodiment, normalization is performed subsequent to correcting for glare and chromatic artifacts.

[0083] In the illustrative embodiment, the invention masks glare and chromatic artifacts from images prior to normalization. In the case of whitening data, glare may have a negative impact, since glare is visually similar to the tissue whitening that is the object of the analysis. Chromatic artifacts may have a more limited impact on the intensity of pixels and may be removed with the temporal and spatial filters described above.

[0084] Thresholding may be used to mask out glare and chromatic artifacts. In the illustrative embodiment thresholding is performed in the $L^*u^*v^*$ color space. Preferably, the invention also employs a correlate for hue, expressed as in Equation (5):

$$h^* = \operatorname{atan}\left(\frac{v^*}{u^*}\right) \tag{5}$$

Since the hue h^* is a periodic function, the illustrative methods of the invention rotate the u^*-v^* plane such that the typical reddish color of the cervix correlates to higher values of h^*. This makes it possible to work with a single threshold for chromatic artifacts. The rotation is given by Equation (6):

$$\begin{pmatrix} u^* \\ v^* \end{pmatrix} \mapsto \left(\begin{matrix} \cos\left(-\frac{\pi}{3}\right) & \sin\left(\frac{\pi}{3}\right) \\ -\sin\left(\frac{\pi}{3}\right) & \cos\left(\frac{\pi}{3}\right) \end{matrix} \right) \begin{pmatrix} u^* \\ v^* \end{pmatrix} . \tag{6}$$

The masks for glare and chromatic artifacts are then respectively obtained using Equations (7) and (8):

$$mask_{\text{glare}} = L^* > 90 \tag{7}$$

$$mask_{\text{hue}} = h^* < \frac{\pi}{5} \quad , \tag{8}$$

where $L^* \in [0,100]$ and $h^* \in [0,2\pi]$. According to the illustrative embodiment, the masks are eroded to create a safety margin, such that they are slightly larger than the corrupted areas.

[0085] Figures 4A and 4B relate to part of step 104 of Figure 1, preprocessing the images. Figure 4A depicts a "maximum" RGB image representation 402 used in preprocessing a sequence of images of tissue according to an illustrative embodiment of the invention. In the illustrative embodiment, the maximum RGB image is computed by taking for each pixel the maximum RGB values in the whole sequence of images.

[0086] Figure 4B depicts the image representation of Figure 4A after applying a manual mask, used in preprocessing a sequence of images of tissue according to an illustrative embodiment of the invention. According to the illustrative embodiment, the method applies the manual mask in addition to the masks for glare and chromatic effects in order to account for obstructions such as hair, foam from the chemical agent, or other obstruction, and/or to narrow analysis to an area of interest. Area 406 of the frame 404 of Figure 4B has been manually masked in accord with the methods of the embodiment.

[0087] Figure 4C, which depicts the image representation of Figure 4B after accounting for glare, is used in preprocessing a sequence of images of tissue according to an illustrative embodiment of the invention. Note that the areas 408, 410, 412, 414, 416, and 418 of the frame 405 of Figure 4C have been masked for glare using Equation (7).

[0088] Figure 4D, which depicts the image representation of Figure 4C after accounting for chromatic artifacts, is used in preprocessing a sequence of images of tissue according to an illustrative embodiment of the invention. The areas 420 and 422 of the frame 407 of Figure 4D have been masked for chromatic artifacts using Equation (8).

[0089] To reduce the amount of data to process and to improve the signal-to-noise ratio of the signals used in the segmentation techniques discussed below, illustrative methods of the invention pre-segment the image plane into *grains*. Illustratively, the mean grain surface is about 30 pixels. However, in other embodiments, it is between about a few pixels and about a few hundred pixels. The segmentation methods can be applied starting at either the pixel level or the grain level.

[0090] One way to "pre-segment" the image plane into grains is to segment each of the images in the sequence using a watershed transform. One goal of the watershed technique is to simplify a gray-level image by viewing it as a three-dimensional surface and by progressively "flooding" the surface from below through "holes" in the surface. In one embodiment, the third dimension is the gradient of an intensity signal over the image plane (further discussed herein below). A "hole" is located at each minimum of the surface, and areas are progressively flooded as the "water level" reaches each minimum. The flooded minima are called catchment basins, and the borders between neighboring catchment basins are called watersheds. The catchment basins determine the pre-segmented image.

[0091] Image segmentation with the watershed transform is preferably performed on the image gradient. If the watershed transform is performed on the image itself, and not the gradient, the watershed transform may obliterate important distinctions in the images. Determination of a gradient image is discussed herein below.

[0092] Segmentation is a process by which an image is split into different regions according to one or more pre-defined criteria. In certain embodiments of the invention, segmentation methods are performed using information from an entire sequence of

images, not just from one image at a time. The area depicted in the sequence of images is split into regions based on a measure of similarity of the detected changes those regions undergo throughout the sequence.

[0093] Segmentation is useful in the analysis of a sequence of images such as in aceto-whitening cervical testing. In an illustrative embodiment, since the analysis of time-series data with a one-pixel resolution is not possible unless motion, artifacts, and noise are absent or can be precisely identified, segmentation is needed. Often, filtering and masking procedures are insufficient to adequately relate regions of an image based on the similarity of the signals those regions produce over a sequence of images. Therefore, the illustrative methods of the invention average time-series data over regions made up of pixels whose signals display similar behavior over time.

[0094] In the illustrative embodiment, regions of an image are segmented based at least in part upon a measure of similarity of the detected changes those regions undergo. Since a measure of similarity between regions depends on the way regions are defined, and since regions are defined based upon criteria involving the measure of similarity, the illustrative embodiment of the invention employs an iterative process for segmentation of an area into regions. In some embodiments, segmentation begins by assuming each pixel or each grain (as determined above) represents a region. These individual pixels or grains are then grouped into regions according to criteria defined by the segmentation method. These regions are then merged together to form new, larger regions, again according to criteria defined by the segmentation method.

[0095] A problem that arises when processing raw image data is its high dimension. With a typical whitening signal for a single pixel described by, for example, a sixty-or-

more-dimensional vector, it is often necessary to reduce data dimensionality prior to processing. In the illustrative embodiment, the invention obtains a scalar that quantifies a leading characteristic of two vectors. More particularly, illustrative methods of the invention take the N-dimensional inner (dot) product of two vectors corresponding to two pixel coordinates. A fitting function based on this dot product is shown in Equation (9). This fitting function quantifies the similarity between the signals at two locations.

$$\varphi(\mathbf{x}_1,\mathbf{x}_2)=\frac{\langle \mathbf{w}(\mathbf{x}_1;t),\mathbf{w}(\mathbf{x}_2;t)\rangle}{\max(\Omega(\mathbf{x}_1),\Omega(\mathbf{x}_2))}\quad, \tag{9}$$

where $\mathbf{x}_1$ and $\mathbf{x}_2$ are two pixel coordinates, and $\Omega(\mathbf{x}_1)=\langle \mathbf{w}(\mathbf{x}_1;t),\mathbf{w}(\mathbf{x}_1;t)\rangle$ is the energy of the signal at location $\mathbf{x}_1$.

[0096] Figure 5 relates to step 108 of Figure 1, determining a measure of similarity between regions in each of a series of images . Figure 5 shows a graph 502 illustrating the determination of a measure of similarity of a time series of mean signal intensity 504 for each of two regions *k* and *l* according to an illustrative embodiment of the invention. Figure 5 represents one step in an illustrative segmentation method in which the similarity between the time-series signals of two neighboring regions is compared against criteria to determine whether those regions should be merged together. The type of measure of similarity chosen may vary depending on the segmentation method employed.

[0097] Curve 506 in Figure 5 represents the mean signal intensity 504 of region *k* in each of a sequence of images and is graphed versus time 505. Mean signal intensity 504 of region *k* is expressed as in Equation (10):

$$\mathbf{w}(k;t) \;=\frac{1}{N_k}\sum_{(i,j)\in\mathbb{S}_k}\mathbf{w}(i,j;t)\quad, \tag{10}$$

where $\mathbb{S}_k\subset\mathbb{N}^2$ is the set of all pixels that belong to the k$^{\text{th}}$ region and N_k is the size of $\mathbb{S}_k$.

[0098] Curve 508 of Figure 5 represents the mean signal intensity 504 of region l and is graphed versus time 505. Mean signal intensity 504 of region l is expressed as in Equation (10), replacing "k" with "l" in appropriate locations. The shaded area 510 of Figure 5 represents dissimilarity between mean signals over region k and region l. The chosen measure of similarity, φ_{kl}, also referred to herein as the fitting function, between regions k and l may depend on the segmentation method employed. For the region merging segmentation technique, discussed in more detail below, as well as for other segmentation techniques, the measure of similarity used is shown in Equation (11):

$$\varphi_{kl} = \frac{\langle \mathbf{w}(k;t), \mathbf{w}(l;t) \rangle}{\max(\Omega(k), \Omega(l))}, \qquad (11)$$

where the numerator represents the N-dimensional dot product of the background-subtracted mean signal intensities 504 of region k and region l; the denominator represents the greater of the energies of the signals corresponding to regions k and l, $\Omega(k)$ and $\Omega(l)$; and $-1 \leq \varphi_{kl} \leq 1$. In this embodiment, the numerator of Equation (11) is normalized by the higher signal energy and not by the square root of the product of both energies.

[0099] In the case of whitening signals, for example, the fitting function defined by Equation (9) can be used to obtain a gradient image representing the variation of whitening values in x-y space. The gradient of an image made up of intensity signals is the approximation of the amplitude of the local gradient of signal intensity at every pixel location. The watershed transform is then applied to the gradient image. This may be done when pre-segmenting images into "grains" as discussed above, as well as when performing the hierarchical watershed segmentation approach and combined method segmentation approaches discussed below.

[0100] A gradient image representing a sequence of images is calculated for an individual image by computing the fitting value $\varphi(i,j;i_o,j_o)$ between a pixel (i_o, j_o) and all its neighbors $(i,j) \in \mathbb{N}_{(i_0,j_0)}$, where $\mathbb{N}_{(i_0,j_0)} = \{(i_0-1,j_0),(i_0,j_0-1),(i_0+1,j_0),(i_0,j_0+1)\}$.

[0101] Since the best fit corresponds to a null gradient, the derivative of the fitting value is computed as in Equation (12):

$$\varphi_\nabla(i,j;i_0,j_0) = \delta\frac{1-\varphi(i,j;i_0,j_0)}{1+\varphi(i,j;i_0,j_0)}, \tag{12}$$

where $\varphi_\nabla(i,j;i_0,j_0) \in (-\infty,\infty)$. The sign of $\varphi_\nabla(i,j;i_0,j_0)$ is given by Equation (13):

$$\delta = \begin{cases} 1 & \text{if } \Omega(i_0,j_0) \geq \Omega(i,j) \\ -1 & \text{if } \Omega(i_0,j_0) < \Omega(i,j) \end{cases}. \tag{13}$$

[0102] Then, the derivatives of the signals are approximated as the mean of the forward and backward differences shown in Equations (14) and (15).

$$\frac{d}{dx}\mathbf{w}(i_0,j_0) = \frac{\varphi_\nabla(i_0,j_0-1;i_0,j_0)-\varphi_\nabla(i_0,j_0+1;i_0,j_0)}{2} \tag{14}$$

$$\frac{d}{dy}\mathbf{w}(i_0,j_0) = \frac{\varphi_\nabla(i_0-1,j_0;i_0,j_0)-\varphi_\nabla(i_0+1,j_0;i_0,j_0)}{2}. \tag{15}$$

The norm of the gradient vector is then calculated from the approximations of Equations (14) and (15).

[0103] Since the fitting values include information from the entire sequence of images, one may obtain a gradient image which includes information from the entire sequence of images, and which, therefore, shows details not visible in all of the images. The gradient image may be used in the watershed pre-segmentation technique discussed herein above and the hierarchical watershed technique discussed herein below. Had the gradient image been obtained from a single reference image, less detail would be included, and the application of a watershed segmentation method to the gradient image would segment the image plane based on less data. However, by using a gradient image as determined from

Equations (14) and (15), the invention enables a watershed segmentation technique to be applied which divides an image plane into regions based on an entire sequence of data, not just a single reference image.

[0104] Figure 6 relates to step 110 of Figure 1, segmenting the area represented in a sequence of images into regions based on measures of similarity between regions over the sequence. Various techniques may be used to perform the segmentation step 110 of Figure 1. Figure 6 shows a schematic flow diagram 602 depicting a region merging technique of segmentation according to an illustrative embodiment of the invention. In this technique, each grain or pixel is initially a region, and the method merges neighboring according to a predefined criterion in an iterative fashion. The criterion is based on a measure of similarity, also called a fitting function. The segmentation converges to the final result when no pair of neighboring regions satisfies the merging criterion. In the case of aceto-whitening data, for instance, it is desired to merge regions whose whitening data is similar. The fitting function will therefore quantify the similarity over the sequence between two signals corresponding to two neighboring regions.

[0105] Thus, the segmentation begins at step 604 of Figure 6, computing the fitting function to obtain "fitting values" for all pairs of neighboring regions. In the region merging approach, this is the measure of similarity provided by Equation (11), where mean signal intensity of a region k is defined by Equation (10). This fitting function is equivalent to the minimum normalized Euclidean distance, δ^2_{kl}, between the mean signal intensities of regions k and l shown in Equation (16):

$$\delta_{kl}^2 = \frac{\left\|\mathbf{w}(k;t)-\mathbf{w}(l;t)\right\|_{L_2}^2}{\max\left(\Omega(k),\Omega(l)\right)} = 2\left(\frac{1}{2}\left(1+\frac{\min\left(\Omega(k),\Omega(l)\right)}{\max\left(\Omega(k),\Omega(l)\right)}\right)-\frac{\left\langle \mathbf{w}(k;t),\mathbf{w}(l;t)\right\rangle}{\max\left(\Omega(k),\Omega(l)\right)}\right). \tag{16}$$

This notation reveals the effect of normalizing using the higher energy of the two signals instead of normalizing each signal by its L_2 norm. The method using Equation (16) or Equation (11) applies an additional "penalty" when both signals have different energies, and therefore, fitting values are below 1.0 when the scaled versions of the two signals are the same, but their absolute values are different.

[0106] In step 606 of Figure 6, the fitting values (measures of similarity) corresponding to pairs of neighboring regions that are larger than a given threshold are sorted from greatest to least. In step 608, sorted pairs are merged according to best fit, keeping in mind that each region can only be merged once during one iteration. For instance, if neighboring regions *k* and *l* have a fitting value of 0.80 and neighboring regions *k* and *m* have a fitting value of 0.79, regions *k* and *l* are merged together, not *k* and *m*. However, region *m* may be merged with another of its neighboring regions during this iteration, depending on the fitting values computed.

[0107] In step 609, the method recalculates fitting values for all pairs of neighboring regions containing an updated (newly merged) region. In the embodiment, Fitting values are not recalculated for pairs of neighboring regions whose regions are unchanged.

[0108] In step 610 of Figure 6, it is determined whether the fitting values of all pairs of neighboring regions are below a given threshold. The fitting function is a measure of similarity between regions; thus, the higher the threshold, the more similar regions have to be in order to be merged, resulting in fewer iterations and, therefore, more regions that are ultimately defined. If the fitting values of all the regions are below the given

threshold, the merging is complete 612. If not, the process beginning at step 606 repeats, and fitting values of the pairs of neighboring regions as newly defined are sorted.

[0109] Once merging is complete 612, a size rule is applied that forces each region whose size is below a given value to be merged with its best fitting neighboring region, even though the fitting value is lower than the threshold. In this way, very small regions not larger than a few pixels are avoided.

[0110] The segmentation method of Figure 6, or any of the other segmentation methods discussed herein, is performed, for example, where each pixel has up to four neighbors: above, below, left, and right. However, in other illustrative embodiments, segmentation is performed where each pixel can has up to eight neighbors or more, which includes diagonal pixels. It should also be noted that images in a given sequence may be sub-sampled to reduce computation time. For instance, a sequence of 100 images may be reduced to 50 images by eliminating every other image from the sequence to be segmented.

[0111] Figures 7A and 7B illustrate step 112 of Figure 1, relating images after segmentation. Figure 7A depicts a segmentation mask 702 produced using the region merging segmentation technique discussed above for an exemplary aceto-whitening sequence. In an illustrative embodiment, the threshold used in step 610 of Figure 6 to produce this segmentation mask 702 is 0.7. Each region has a different label, and is represented by a different color in the mask 702 in order to improve the contrast between neighboring regions. Other illustrative embodiments use other kinds of display techniques known in the art, in order to relate images of the sequence based on segmentation and/or based on the measure of similarity.

[0112] Figure 7B shows a graph 750 depicting mean signal intensities 752 of segmented regions represented in Figure 7A as functions of a time index 754 according to an illustrative embodiment of the invention. The color of each data series in Figure 7B corresponds to the same-colored segment depicted in Figure 7A. This is one way to visually relate a sequence of images using the results of segmentation. For instance, according to the illustrative embodiment, regions having a high initial rate of increase of signal intensity 752 are identified by observing data series 756 and 758 in Figure 7B, whose signal intensities 752 increase more quickly than the other data series. The location of the two regions corresponding to these two data series is found in Figure 7A. In another example, kinetic rate constants are derived from each of the data series determined in Figure 7B, and the regions having data series most closely matching kinetic rate constants of interest are identified. In another example, one or more data series are curve fit to obtain a characterization of the mean signal intensities 752 of each data series as functions of time.

[0113] Mean signal intensity may have a negative value after background subtraction. This is evident, for example, in the first part of data series 760 of Figure 7B. In some examples, this is due to the choice of the reference frame for background subtraction. In other examples, it is due to negative intensities corresponding to regions corrupted by glare that is not completely masked from the analysis.

[0114] Figure 8 relates to step 110 of Figure 1, segmenting the area represented in a sequence of images into regions based on measures of similarity between regions over the sequence. Figure 8 shows a schematic flow diagram 802 depicting a robust region merging approach of segmentation according to an illustrative embodiment of the

invention. One objective of the robust region merging approach is to take into account the “homogeneity” of data inside the different regions. While the region merging approach outlined in Figure 6 relies essentially on the mean signal of each region to decide subsequent merging, the robust region merging approach outlined in Figure 8 controls the maximum variability allowed inside each region. More specifically, the variance signal, $\sigma^2_w(k;t)$, associated with each region, k, is computed as in Equation (17):

$$\begin{aligned}\sigma_{\mathrm{w}}^2(k;t) &= \frac{1}{N_k}\sum_{(i,j)\in S_k}(\mathbf{w}(i,j;t)-\mathbf{w}(k;t))^2 \\ &= \frac{1}{N_k}\sum_{(i,j)\in S_k}\mathbf{w}^2(i,j;t)-\left(\frac{1}{N_k}\sum_{(i,j)\in S_k}\mathbf{w}(i,j;t)\right)^2\end{aligned} \quad , \qquad (17)$$

where $\mathbf{w}(k;t)$ is mean signal intensity of region k as expressed in Equation (10). The merging criterion is then the energy of the standard deviation signal, computed as in Equation (18):

$$\langle\sigma_{\mathrm{w}},\sigma_{\mathrm{w}}\rangle = \sum_{t\in T}\sigma_{\mathrm{w}}^2(k;t) \quad . \qquad (18)$$

[0115] Segmentation using the illustrative robust region merging approach begins at step 804 of the schematic flow diagram 802 of Figure 8. In step 804 of Figure 8, the variance signal, $\sigma^2_w(k;t)$, of Equation (17) is computed for each region k. Then variance signal energy, or the energy of the standard deviation signal as shown in Equation (18), is calculated for each region k. In step 806 of Figure 8, the values of variance signal energy that are larger than a given threshold are sorted. This determines *which* regions can be merged, but not *in which order* the regions may be merged. In step 808 of Figure 8, the sorted pairs are merged according to the increase in variance each merged pair would create, $\Delta_\sigma(k,l)$, given by Equation (19):

$$\Delta_\sigma(k,l) = \sum_{t\in T}\left(\sigma_{\mathrm{w}}^2(k\cup l;t)-\frac{1}{2}\left[\sigma_{\mathrm{w}}^2(k;t)+\sigma_{\mathrm{w}}^2(l;t)\right]\right), \qquad (19)$$

where k and l represent two neighboring regions to be merged. Thus, if a region can merge with more than one of its neighbors, it merges with the one that increases less the variance shown in Equation (19). Another neighbor may merge with the region in the next iteration, given that it still meets the fitting criterion with the updated region.

[0116] According to the illustrative embodiment, it is possible that a large region neighboring a small region will absorb the small region even when the regions have different signals. This results from the illustrative merging criterion being size-dependent, and the change in variance is small if the smaller region is merged into the larger region. According to a further embodiment, the methods of the invention apply an additional criterion as shown in step 807 of Figure 8 prior to merging sorted pairs in step 808. In step 807, fitting values corresponding to the pairs of neighboring regions are checked against a threshold. The fitting values are determined as shown in Equation (11), used in the region-merging approach. According to the illustrative embodiment, a candidate pair of regions are not merged if its fitting value is below the threshold (e.g., if the two regions are too dissimilar). In the illustrative embodiment, the invention employs a fixed similarity criterion threshold of about $\varphi_{kl} = 0.7$. This value is low enough not to become the main criterion, yet the value is high enough to avoid the merging of regions with very different signals. However, other φ_{kl} values may be used without deviating from the scope of the invention.

[0117] In step 809 of Figure 8, values of the variance signal are recalculated for pairs of neighboring regions containing an updated (newly-merged) region. According to an embodiment of the invention, variance signal values are not recalculated for pairs of neighboring regions whose regions are unchanged.

[0118] In step 810 of Figure 8, the illustrative method of the invention determines whether the values of the variance signal energy for all regions are below a given variance threshold. If all values are below the threshold, the merging is complete 812. If not, the process beginning at step 806 is repeated, and values of variance signal energy of neighboring pairs of regions above the threshold are sorted.

[0119] Figures 9A and 9B illustrate step 112 of Figure 1, relating images after segmentation. Figure 9A depicts a segmentation mask 902 produced using the robust region merging segmentation technique discussed above for an exemplary aceto-whitening sequence. In the illustrative embodiment, the variance threshold used in step 810 of Figure 8 to produce the segmentation mask 902 is 70. However, other variance thresholds may be employed without deviating from the scope of the invention. Each region has a different label, and is represented by a different color in the mask 902 to improve the contrast between neighboring regions. Other display techniques may be used to relate images of the sequence based on segmentation and/or based on the measure of similarity.

[0120] Figure 9B shows a graph 950 depicting mean signal intensity 952 of segmented regions represented in Figure 9A as functions of a time index 954 according to an illustrative embodiment of the invention. The color of each curve in Figure 9B corresponds to the same-colored segment depicted in Figure 9A. This is one way to visually relate a sequence of images using the results of segmentation.

[0121] According to the illustrative embodiment, the method observes data series 956, 958, 960, and 962 in Figure 9B, whose signal intensities 952 increase more quickly than the other data series. The location of the four regions corresponding to these four data

series are in Figure 9A. In another embodiment, the method derives kinetic rate constants from each of the data series determined in Figure 9B, and the regions having data series most closely matching kinetic rate constants of interest are identified. In another example, the method curve fits one or more data series to obtain a characterization of the mean signal intensities 952 of each data series as functions of time.

[0122] Figure 10 relates to step 110 of Figure 1, segmenting an area represented in a sequence of images into regions based on measures of similarity between regions over the sequence, according to one embodiment of the invention. Figure 10 shows a schematic flow diagram 1002 depicting a segmentation approach based on the clustering of data, or more precisely, a "fuzzy c-means" clustering segmentation approach, used in an embodiment of the invention. An objective of this clustering approach is to group pixels into clusters with similar values. Unlike the region merging and robust region merging approaches above, the method does not merge regions based on their spatial relation to each other. In other words, two non-neighboring regions may be merged, depending on the criterion used.

[0123] Let $\mathbb{X} = \{\mathbf{x}_1, \ldots, \mathbf{x}_n\} \subset \mathbb{R}^d$ be a set of n d-dimensional vectors. An objective of clustering is to split $\mathbb{X}$ into c subsets, called partitions, that minimize a given functional, J_m. In the case of the fuzzy c-means, this functional is given by Equation (20):

$$J_m = \sum_{k=1}^{n}\sum_{i=1}^{c}(u_{ik})^m \|\mathbf{x}_k - \mathbf{v}_i\|^2 \quad , \tag{20}$$

where $\mathbf{v}_i$ is the "center" of the i^{th} cluster, $u_{ik} \in [0,1]$ is called the *fuzzy membership* of $\mathbf{x}_k$ to $\mathbf{v}_i$, with $\sum_{i=1}^{c} u_{ik} = 1$, and $m \in [1,\infty]$ is a weighting exponent. The inventors have used m=2 in exemplary embodiments. The distance $\|\cdot\|$ is any inner product induced norm on

$\mathbb{R}^d$. The minimization of J_m as defined by Equation (20) leads to the following iterative system:

$$u_{ik} = \left(\sum_{j=1}^{c} \left(\frac{\|\mathbf{x}_k - \mathbf{v}_i\|}{\|\mathbf{x}_k - \mathbf{v}_j\|} \right)^{\frac{2}{m-1}} \right)^{-1}, \tag{21}$$

$$\mathbf{v}_i = \frac{\sum_{k=1}^{n} (u_{ik})^m \mathbf{x}_k}{\sum_{k=1}^{n} (u_{ik})^m}. \tag{22}$$

[0124] The distance, $\|\mathbf{x}_k - \mathbf{v}_i\|$, is based on the fitting function given in Equation (11). If it is assumed that similar signals are at a short distance from each other, then Equation (23) results:

$$\|\mathbf{x}_k - \mathbf{v}_i\| = \frac{1}{2}(1 - \varphi_{ki}) \ , \tag{23}$$

where φ_{ki} is given by Equation (11).

[0125] Thus, in this embodiment, the segmentation begins at step 1004 of Figure 10; and the method initializes values of $\mathbf{v}_i$, where i=1 to c and c is the total number of clusters. In one embodiment, the method sets the initial value of $\mathbf{v}_i$ randomly. In step 1006 of Figure 10, the method calculates values of u_{ik}, the fuzzy membership of $\mathbf{x}_k$ to $\mathbf{v}_i$, according to Equation (21). In step 1008 of Figure 10, the method updates values of $\mathbf{v}_i$ according to Equation (22), using the previously determined value of u_{ik},. The algorithm converges when the relative decrease of the functional J_m as defined by Equation (20) is below a predefined threshold, for instance, 0.001. Thus, in step 1010 of Figure 10, an embodiment of the method determines whether the relative decrease of J_m is below the threshold. If not, then the process beginning at step 1006 is repeated. If the relative decrease of J_m is below the threshold, then the segmentation is completed by labeling each pixel according to its highest fuzzy membership, e.g. $\operatorname{maxind}_{i\in\{1,\ldots,c\}}\{u_{ik}\}$. Some

embodiments have more regions than clusters, since pixels belonging to different regions with similar signals can contribute to the same cluster.

[0126] Figures 11A and 11B show an illustrative embodiment of the method at step 112 of Figure 1. Figure 11A depicts a segmentation mask 1102 produced using the clustering technique discussed above for an exemplary aceto-whitening sequence, according to the embodiment. The number of clusters, c, chosen in this embodiment is 3. There are more regions than clusters, since portions of some clusters are non-contiguous. The threshold for the relative decrease of J_m in step 1010 of Figure 10 is chosen as 0.001 in this embodiment. Each cluster has a different label and is represented by a different color in the mask 1102. Other embodiments employ other kinds of display techniques to relate images of the sequence based on segmentation and/or based on the measure of similarity.

[0127] Figure 11B shows a graph 1120 depicting mean signal intensities 1122 of segmented regions represented in Figure 11A as functions of a time index 1124 according to an illustrative embodiment of the invention. The color of each data series in Figure 11B corresponds to the same-colored cluster depicted in Figure 11A. The graph 1120 of Figure 11B is one way to visually relate a sequence of images using the results of segmentation according to this embodiment. In this embodiment, the method identifies a cluster having a high initial rate of increase of signal intensity 1122 by observing data series 1126 in Figure 11B, whose signal intensity increases more quickly than the other data series. Regions belonging to the same cluster have data series 1126 of the same color. In another embodiment, the method derives kinetic rate constants from each of the data series determined in Figure 11B, and the regions having data series most closely matching kinetic rate constants of interest are identified. In another example, the method

curve fits one or more data series to obtain characterization of the mean signal intensities 1122 of each data series as functions of time.

[0128] Similarly, Figures 11C and 11D illustrate an embodiment of the method at step 112 of Figure 1, relating images after segmentation. Figure 11C depicts a segmentation mask 1140 produced using the clustering technique for the exemplary aceto-whitening sequence in Figure 11A, according to the embodiment. In Figure 11C, however, the number of clusters, c, chosen is 2. Again, there are more regions than clusters, since portions of the same clusters are non-contiguous. The threshold for the relative decrease of J_m in step 1010 of Figure 10 is 0.001 for this embodiment.

[0129] Figure 12 relates to step 110 of Figure 1, segmenting the area represented in a sequence of images into regions based on measures of similarity between regions over the sequence, according to an illustrative embodiment of the invention. In this embodiment, morphological techniques of segmentation are continued beyond filtering and pre-segmentation. Figure 12 shows a schematic flow diagram 1202 depicting a hierarchical watershed approach of segmentation according to the illustrative embodiment.

[0130] In step 1204 of Figure 12, the method computes a gradient image from Equations (14) and (15) according to the embodiment, which incorporates data from the entire sequence of images, as discussed above. The method segments data based on information from the entire sequence of images, not just one image. A watershed transform is applied to this gradient image in step 1206 of Figure 12. Some embodiments apply first-in-first-out (FIFO) queues, sorted data, and other techniques to speed the performance of the watershed transform. Also, in some embodiments, the method

applies a sigmoidal scaling function to the gradient image , prior to performing the watershed transform to enhance the contrast between whitish and reddish (dark) regions, which is particularly useful when analyzing images of a cervix. The catchment basins resulting from application of the watershed transform represent the segmented regions in this embodiment.

[0131] Figure 13 shows a gradient image 1302 calculated from Equations (5) and (6) for an exemplary sequence of images, according to an embodiment of the invention.

[0132] According to an embodiment, the method at step 1208 of Figure 12 constructs a new gradient image using geodesic reconstruction. Two different techniques of geodesic reconstruction which embodiments may employ include erosion and dilation. In step 1210 of Figure 12, the method determines whether over-segmentation has been reduced sufficiently, according to this embodiment. If so, the segmentation may be considered complete, or one or more additional segmentation techniques may be applied, such as a region merging or robust region merging technique, both of which are discussed above. If over-segmentation has not been reduced sufficiently, the method calculates the watershed transform of the reconstructed gradient image as in step 1206, and the process is continued.

[0133] Certain embodiment methods use the hierarchical watershed to segment larger areas, such as large lesions or the background cervix. In some embodiments, the number of iterations is less than about 4 such that regions do not become too large, obscuring real details.

[0134] In some embodiments, the method performs one iteration of the hierarchical watershed, and continues merging regions using the robust region merging technique.

[0135] Figures 14A and 14B show an illustrative embodiment of the method at step 112 of Figure 1, relating images after segmentation. Figure 14A depicts a segmentation mask 1402 produced using one iteration of the hierarchical watershed technique discussed above for an exemplary aceto-whitening sequence, according to the embodiment. Each region has a different label, and is represented by a different color in the mask 1402. Figure 14B depicts a segmentation mask 1430 produced using two iterations of the hierarchical watershed technique discussed above for the exemplary aceto-whitening sequence, according to the embodiment. The segmentation mask 1430 in Figure 14B, produced using two iterations, has fewer regions and is more simplified than the segmentation mask 1402 in Figure 14A, produced using one iteration.

[0136] Other embodiments employ a "region growing technique" to performing step 110 of Figure 1, segmenting an area represented in a sequence of images into regions based on measures of similarity between regions over the sequence. The region growing technique is different from the region merging and robust region merging techniques in that one or more initial regions, called *seed regions*, grow by merging with neighboring regions. The region merging, robust region merging, and region growing methods are each iterative. The region merging and region growing techniques each use the same fitting function to evaluate the similarity between signals from neighboring regions. In some embodiments of the region growing algorithm, the user manually selects seed regions. In other embodiments, the seed regions are selected in an automatic fashion. Hard criteria may be used to select areas that are of high interest and/or which behave in a certain way. In some embodiments, a user selects seed regions based on that user's

experience. The region growing algorithm then proceeds by detecting similar regions and adding them to the set of seeds.

[0137] One embodiment of the invention is a combined technique using the region growing algorithm, starting with a grain image, followed by performing one iteration of the hierarchical watershed technique, and then growing the selected region according to the robust region merging algorithm.

[0138] In some embodiments, the segmentation techniques discussed herein are combined in various ways. In some embodiments, the method processes and analyzes data from a sequence of images in an aceto-whitening test, for instance, using a coarse-to-fine approach. In one embodiment, a first segmentation reveals large whitening regions, called background lesions, which are then considered as regions of interest and are masked for additional segmentation.

[0139] A second segmentation step of the embodiment may outline smaller regions, called foreground lesions. Segmentation steps subsequent to the second step may also be considered. As used here, the term "lesion" does not necessarily refer to any diagnosed area, but to an area of interest, such as an area displaying a certain whitening characteristic during the sequence. From the final segmentation, regions are selected for diagnosis, preliminary or otherwise; for further analysis; or for biopsy, for example. Additionally, the segmentation information may be combined with manually drawn biopsy locations for which a pathology report may be generated. In one illustrative embodiment, the method still applies the pre-processing procedures discussed herein above before performing the multi-step segmentation techniques.

[0140] Figure 16A shows a segmentation mask produced using a combined clustering approach and robust region merging approach for an exemplary aceto-whitening sequence, according to an illustrative embodiment of the invention. In the embodiment, the method performs pre-processing steps, including pre-segmenting pixels into grains using a watershed transform as discussed herein above. Then, the method applies the clustering technique to the sequence as discussed in Figure 10, using c = 3 clusters and $J_m = 0.001$. This produces a "coarse" segmentation. From this coarse segmentation, the method selects a boomerang-shaped background lesion, corresponding to a large whitening region. The method masks out, or eliminates from further analysis, the remaining areas of the image frame.

[0141] Then, a robust region merging procedure is applied, as shown in Figure 8, to the background lesion, according to the embodiment. Here, the method uses a similarity criterion of $\varphi_{kl} = 0.7$ in step 807 of Figure 8, and a variance threshold of 120 in step 810 of Figure 8. In this and other embodiments, regions less than 16 pixels large are removed. The resulting segmentation is shown in frame 1602 of Figure 16A.

[0142] Figure 16B shows a graph 1604 depicting mean signal intensity 1606 of segmented regions represented in Figure 16A as functions of a time index 1608 according to an illustrative embodiment of the invention. The color of each curve in Figure 16B corresponds to the same-colored segment depicted in Figure 16A. Regions having a high initial rate of increase of signal intensity 1606 include regions 1610 and 1612, shown in Figure 16A.

[0143] Figure 17A represents a segmentation mask produced using a combined clustering approach and watershed approach for the exemplary aceto-whitening sequence

of Figure 16A, according to an illustrative embodiment of the invention. The method performs pre-processing steps, including the pre-segmenting of pixels into grains using a watershed transform as discussed herein above. Then, the method applies the clustering technique to the sequence as discussed in Figure 10, using c = 3 clusters and $J_m = 0.001$. This produces a "coarse" segmentation. From this coarse segmentation, the method selects a boomerang-shaped background lesion, corresponding to a large whitening region. The method masks out remaining areas of the image frame.

[0144] Then, the method applies a hierarchical watershed segmentation procedure, as shown in Figure 12. In this embodiment, the method computes one iteration of the watershed transform, then a region merging technique as per Figure 6, using a fitting value threshold of 0.85 in step 610 of Figure 6. Regions smaller than 16 pixels are removed. The resulting segmentation is shown in frame 1702 of Figure 17A.

[0145] Figure 17B shows a graph 1720 depicting mean signal intensity 1722 of segmented regions represented in Figure 17A as functions of a time index 1724 according to an illustrative embodiment of the invention. The color of each curve in Figure 17B corresponds to the same-colored segment depicted in Figure 17A. Regions having a high initial rate of increase of signal intensity 1722 include regions 1726 and 1728, shown in Figure 17A.

[0146] Figure 18A represents a segmentation mask produced using a two-step clustering approach for the exemplary aceto-whitening sequence of Figure 16A, according to an illustrative embodiment of the invention. The method performs pre-processing steps, including pre-segmenting pixels into grains using a watershed transform as discussed herein above. Then, the method applies a clustering technique to the

sequence as discussed in Figure 10, using c = 3 clusters and $J_m = 0.001$. This produces a "coarse" segmentation. From this coarse segmentation, the method selects a boomerang-shaped background lesion, corresponding to a large whitening region. The method masks out the remaining areas of the image frame from further analysis.

[0147] Then, the method applies a second clustering procedure, as shown in Figure 10, to the background lesion. Here again, the method uses c = 3 clusters and $J_m = 0.001$. Regions less than 16 pixels large are removed. This produces a foreground lesion, shown in frame 1802 of Figure 18A.

[0148] Figure 18B shows a graph 1820 depicting mean signal intensity 1822 of segmented regions represented in Figure 18A as functions of a time index 1824 according to an illustrative embodiment of the invention. The color of each curve in Figure 18B corresponds to the same-colored cluster depicted in Figure 18A. Regions having a high initial rate of increase of signal intensity 1822 include regions 1828 and 1826, shown in Figure 18A.

[0149] Figure 19 depicts the human cervix tissue of Figure 2A with an overlay of manual doctor annotations made after viewing the exemplary aceto-whitening image sequence discussed herein above. Based on her viewing of the sequence and on her experience with the aceto-whitening procedure, the doctor annotated regions with suspicion of pathology 1904, 1906, 1908, 1910, and 1912. The doctor did not examine results of any segmentation analysis prior to making the annotations. Regions 1910 and 1912 were singled out by the doctor as regions with the highest suspicion of pathology.

[0150] Figure 20A is a representation of a segmentation mask produced using a combined clustering approach and robust region merging approach as discussed above

and as shown in Figure 16A, according to an illustrative embodiment of the invention. The segmentation mask in Figure 20A, however, is shown with a correspondingly-aligned overlay of the manual doctor annotations of Figure 19. Figure 20B is a representation of a segmentation mask produced using a combined clustering approach and watershed technique as discussed above and shown in Figure 18A. The segmentation mask in Figure 20B, however, is shown with a correspondingly-aligned overlay of the manual doctor annotations of Figure 19, according to an embodiment of the invention.

[0151] Areas 1912 and 1910 in Figures 20A and 20B correspond to the doctor's annotations of areas of high suspicion of pathology. In the segmentation masks produced from the combined techniques of both Figures 20A and 20B, these areas (1912 and 1910) correspond to regions of rapid, intense whitening. In the doctor's experience, areas of rapid, intense whitening correspond to areas of suspicion of pathology. Thus, the techniques discussed herein provide a method of determining a tissue characteristic, namely, the presence or absence of a suspicion of pathology. Certain embodiments of the invention use the techniques in addition to a doctor's analysis or in place of a doctor's analysis. Certain embodiments use combinations of the methods described herein to produce similar results.

[0152] Some embodiments of the invention for applications other than the analysis of acetowhitening tests of cervical tissue also use various inventive analysis techniques as described herein. A practitioner may customize elements of an analysis technique disclosed herein, based on the attributes of her particular application, according to embodiments of the invention. For instance, the practitioner may choose among the segmentation techniques disclosed herein, depending on the application for which she

intends to practice embodiments of the inventive methods. By using the techniques described herein, it is possible to visually capture all the frames of a sequence at once and relate regions according to their signals over a period of time.

[0153] Certain embodiments of the invention methods analyze more complex behavior. Some embodiments segment the image plane of a sequence of images, then feature-extract the resulting mean intensity signals to characterize the signals of each segmented region. Examples of feature extraction procedures include any number of curve fitting techniques or functional analysis techniques used to mathematically and/or statistically describe characteristics of one or more data series. In some embodiments, these features are then used in a manual, automated, or semi-automated method for the classification of tissue.

[0154] For example, in certain embodiments, the method classifies a region of cervical tissue either as "high grade disease" tissue, which includes Cervical Intraepithelial Neoplasia II/III (CIN II/III), or as "not high grade disease" tissue, which includes normal squamous (NED – no evidence of disease), metaplasia, and CIN I tissue. The classification for a segmented region may be within a predicted degree of certainty using features extracted from the mean signal intensity curve corresponding to the region. In one embodiment, this classification is performed for each segmented region in an image plane to produce a map of regions of tissue classified as high grade disease tissue. Other embodiments make more specific classifications and distinctions between tissue characteristics, such as distinction between NED, metaplasia, and CIN I tissue.

[0155] Figure 21A, Figure 21B, Figure 21C, and Figure 21D depict steps in the classification of regions of tissue in a sequence of images obtained during an

acetowhitening procedure performed on a patient with high grade disease according to an illustrative embodiment of the invention. Figure 21A depicts a reference image 2102 of cervical tissue of the patient from a sequence of images obtained during the acetowhitening test. Figure 21B is a representation 2106 of the reference image 2102 of Figure 21A after applying a manual mask, accounting for glare, and accounting for chromatic artifacts as discussed herein according to an illustrative embodiment of the invention. For example, areas such as areas 2110 and 2112 of Figure 21B have been masked for glare and chromatic effects, respectively, using techniques as discussed herein. Figure 21C shows a graph 2120 depicting mean signal intensities 2122 of segmented regions for the sequence of Figure 21A as functions of a time index 2124 and as determined using the clustering segmentation approach depicted in the schematic flow diagram 1002 of Figure 10 and as discussed herein according to an illustrative embodiment of the invention.

[0156] It was desired to classify each of the segmented regions as either "indicative of high grade disease" or "not indicative of high grade disease." Thus, an embodiment of the invention extracted specific features from each of the mean signal intensity data series depicted in the graph 2120 of Figure 21C, and used these features in a classification algorithm.

[0157] A classification algorithm was designed using results of a clinical study. In the study, mean signal intensity curves were determined using sequences of images from acetowhitening tests performed on over 200 patients. The classification algorithm may be updated according to an embodiment of the invention upon conducting further or different clinical testing. The present algorithm was based upon two feature parameters

extracted from each of certain mean signal intensity data series corresponding to segmented regions of the image sequences for which biopsies were performed. These two feature parameters are as follows:

1. X is the slope of a curve (here, a polynomial curve) fitted to the mean signal intensity data series of a segmented region at the time corresponding to 235 seconds after application of the acetic acid (M235); and

2. Y is the slope of the polynomial curve at the time corresponding to an intensity that is –16dB from the maximum intensity (about 45% of the maximum mean signal intensity) on the decaying side of the polynomial curve (-16dB slope).

The choice of the feature parameters X and Y above was made by conducting a Discrimination Function (DF) analysis of the data sets from the clinical study. A wide range of candidate feature parameters, including X and Y, were tested. X and Y provided a classification algorithm having the best accuracy.

[0158] A jackknifed classification matrix linear discriminant analysis was performed on the extracted features X and Y corresponding to certain of the mean signal intensity curves from each of the clinical tests. The curves used were those corresponding to regions for which tissue biopsies were performed. From the linear discriminant analysis, it was determined that a classification algorithm using the discriminant line shown in Equation (24) results in a diagnostic sensitivity of 88% and a specificity of 88% for the separation of CIN II/III (high grade disease) from the group consisting of normal squamous (NED), metaplasia, and CIN I tissue (not high grade disease):

$$Y = -0.9282X - 0.1348 \ . \qquad (24)$$

Varying the classification model parameters by as much as 10% yields very similar model outcomes, suggesting the model features are highly stable.

[0159] Figure 21D represents a map 2130 of regions of tissue as segmented in Figure 21C classified as either high grade disease tissue or not high grade disease tissue using the classification algorithm of Equation (24). This embodiment determined this classification for each of the segmented regions by calculating X and Y for each region and determining whether the point (X,Y) falls below the line of Equation (24), in which case the region was classified as high grade disease, or whether the point (X,Y) falls above the line of Equation (24), in which case the region was classified as not high grade disease. The embodiment draws further distinction depending on how far above or below the line of Equation (24) the point (X,Y) falls. The map 2130 of Figure 21D indicates segmented regions of high grade disease as red, orange, and yellow, and regions not classifiable as high grade disease as blue. The index 2132 reflects how far the point (X,Y) of a given segment falls below the line of Equation (24). Other embodiments include those employing other segmentation techniques as described herein. Still other embodiments include those employing different classification algorithms, including those using feature parameters other than X and Y, extracted from mean signal data series corresponding to segmented regions.

Equivalents

[0160] While the invention has been particularly shown and described with reference to specific preferred embodiments, it should be understood by those skilled in the art that various changes in form and detail may be made therein without departing from the spirit and scope of the invention as defined by the appended claims.

What is claimed is: